Appendix A: Example Generative AI–Trustworthy Characteristic Crosswalk

A.1: Trustworthy Characteristic to Generative AI Risk Crosswalk

Table A.1: Trustworthy Characteristic to Generative AI Risk Crosswalk.

Accountable and Transparent	**Explainable and Interpretable**	**Fair with Harmful Bias Managed**	**Privacy Enhanced**
Data Privacy	Human-AI Configuration	Confabulation	Data Privacy
Environmental	Value Chain and Component Integration	Environmental	Human-AI Configuration
Human-AI Configuration		Human-AI Configuration	Information Security
Information Integrity		Intellectual Property	Intellectual Property
Intellectual Property		Obscene, Degrading, and/or Abusive Content	Value Chain and Component Integration
Value Chain and Component Integration		Toxicity, Bias, and Homogenization	
		Value Chain and Component Integration	

Safe	**Secure and Resilient**	**Valid and Reliable**
CBRN Information	Dangerous or Violent Recommendations	Confabulation
Confabulation	Data Privacy	Human-AI Configuration
Dangerous or Violent Recommendations	Human-AI Configuration	Information Integrity
Data Privacy	Information Security	Information Security
Environmental	Value Chain and Component Integration	Toxicity, Bias, and Homogenization
Human-AI Configuration		Value Chain and Component Integration
Information Integrity		
Information Security		
Obscene, Degrading, and/or Abusive Content		
Value Chain and Component Integration		

Usage Note: Table A.1 provides an example of mapping GAI risks onto AI RMF trustworthy characteristics. Mapping GAI risks to AI RMF trustworthy characteristics can be particularly useful when existing policies, processes, or controls can be applied to manage GAI risks, but have been previously implemented in alignment with the AI RMF trustworthy characteristics. Many mappings are possible. Mappings that differ from the example may be more appropriate to meet a particular organization's risk management goals.

A.2: Generative AI Risk to Trustworthy Characteristic Crosswalk

Table A.2: Generative AI Risk to Trustworthy Characteristic Crosswalk.

CBRN Information	Confabulation	Dangerous or Violent Recommendations	Data Privacy
Safe	Fair with Harmful Bias Managed Safe Valid and Reliable	Safe Secure and Resilient	Accountable and Transparent Privacy Enhanced Safe Secure and Resilient

Environmental	Human-AI Configuration	Information Integrity	Information Security
Accountable and Transparent Fair with Harmful Bias Managed Safe	Accountable and Transparent Explainable and Interpretable Fair with Harmful Bias Managed Privacy Enhanced Safe Secure and Resilient Valid and Reliable	Accountable and Transparent Safe Valid and Reliable	Privacy Enhanced Safe Secure and Resilient Valid and Reliable

Intellectual Property	Obscene, Degrading, and/or Abusive Content	Toxicity, Bias, and Homogenization	Value Chain and Component Integration
Accountable and Transparent Fair with Harmful Bias Managed Privacy Enhanced	Fair with Harmful Bias Managed Safe	Fair with Harmful Bias Managed Valid and Reliable	Accountable and Transparent Explainable and Interpretable Fair with Harmful Bias Managed Privacy Enhanced Safe Secure and Resilient Valid and Reliable

Usage Note: Table A.2 provides an example of mapping AI RMF trustworthy characteristics onto GAI risks. Mapping AI RMF trustworthy characteristics to GAI risks can assist organizations in aligning GAI guidance to existing AI/ML policies, processes, or controls or to extend GAI guidance to address additional AI/ML technologies. Many mappings are possible. Mappings that differ from the example may be more appropriate to meet a particular organization's risk management goals.

A.3: Traditional Banking Risks, Generative AI Risks and Trustworthy Characteristics Crosswalk

Table A.3: Traditional Banking Risks, Generative AI Risks and Trustworthy Characteristics Crosswalk.

Compliance Risk	**Information Security Risk**	**Legal Risk**	**Model Risk**
Data Privacy Information Security Toxicity, Bias, and Homogenization Value Chain and Component Integration	Data Privacy Information Security Value Chain and Component Integration	Intellectual Property Obscene, Degrading, and/or Abusive Content Value Chain and Component Integration	Confabulation Dangerous or Violent Recommendations Information Integrity Obscene, Degrading, and/or Abusive Content Toxicity, Bias, and Homogenization
Accountable and Transparent Fair with Harmful Bias Managed Privacy Enhanced Secure and Resilient	Privacy Enhanced Secure and Resilient	Accountable and Transparent Safe	Valid and Reliable

Operational Risk	**Reputational Risk**	**Strategic Risk**	**Third Party Risk**
Confabulation Human-AI Configuration Information Security Value Chain and Component Integration	Confabulation Dangerous or Violent Recommendations Environmental Human-AI Configuration Information Integrity Obscene, Degrading, and/or Abusive Content Toxicity, Bias, and Homogenization	Environmental Information Integrity Information Security Value Chain and Component Integration	Information Integrity Value Chain and Component Integration
Safe Secure and Resilient Valid and Reliable	Accountable and Transparent Fair with Harmful Bias Managed Valid and Reliable	Accountable and Transparent Secure and Resilient Valid and Reliable	Accountable and Transparent Explainable and Interpretable

Usage Note: Table A.3 provides an example of mapping GAI risks and AI RMF trustworthy characteristics. This type of mapping can enable incorporation of new AI guidance into existing policies, processes, or controls or the application of existing policies, processes, or controls to newer AI risks.